

INFORMATION STRUCTURING IN PAPAGO NARRATIVE DISCOURSE

DORIS L. PAYNE

*Summer Institute of Linguistics
and University of Oregon*

In previous accounts, the word order of Papago (Uto-Aztecan, Arizona) has been described—in terms of syntactic roles—as SOV, SVO, and VSO. However, discourse data show that (surface) order is most insightfully accounted for by strong pragmatic principles. Information preceding the verb is either pragmatically marked, or is information for which the hearer is instructed to 'open a new active discourse file'. Other information follows the verb. Syntactic role correlates highly with the discourse-pragmatic and semantic status of the information encoded, but has almost no correlation with order. Therefore, in terms of the typological tradition represented by J. Greenberg, J. Hawkins, and others, no particular order of syntactic roles should be taken as basic. Rather, languages in which order is not based on syntactic role should simply not be forced into an order typology based on syntactic role.*

1. INTRODUCTION. Identifying the so-called 'basic word order' of subject, object, and verb for any given language is not without difficulties. First, most writers who deal with this topic are not concerned with order of words or constituents, but rather with order of the syntactic-semantic roles subject and object—or possibly of S, A, and O in the sense of Comrie 1978, Dixon 1979, and Silverstein 1976.¹ Within the typological tradition largely initiated by Greenberg 1966, basic order is often identified (partly) in terms of what is least pragmatically marked, most frequent in natural discourse, or most consistent throughout the grammar of the language (Hawkins 1983, Mallinson & Blake 1981). To acknowledge the seminal role of Greenberg's work, and for ease of reference, this is hereafter termed the 'standard typological tradition'.

A second tradition assumes that the term 'basic order' should be reserved for that underlying or deep order of roles from which sentence-level syntactic facts of the language can be best predicted, regardless of whether this is the most consistent or least marked order in natural discourse. In fact, the relevance of discourse data for determining underlying order is sometimes explic-

* I would like to thank Ethelene Rosero and Dean Saxton, who have generously helped me learn something about Pima and Papago. In addition, William Bright, Wallace Chafe, Scott DeLancey, John Du Bois, Kenneth Hale, Talmy Givón, Pamela Munro, Tom Payne, and Sandra Thompson have given helpful comments on a previous version of this paper. Much appreciation goes to Nonie Holtz for consultation on the statistical analyses.

¹ S is defined as the only argument of a single argument clause (cf. Comrie 1978, Dixon 1979). A is the most agent-like participant of a multi-argument clause, or that entity which is morpho-syntactically treated as an agent would most commonly be; it is generally that argument which would be the subject in a 'non-ergative' language. O (Comrie uses P) is the next most agent-like participant of a multi-argument clause (in fact, it is often not agent-like at all); it is that argument which would be the object in a 'non-ergative' language. When S, A, and O—or subject and object—are encoded by full noun phrases or free pronouns, the encoding device usually forms a single syntactic constituent. To facilitate discussion, I will continue using the relatively neutral term 'constituent'.

itly rejected: since underlying order is taken as a strictly syntactic fact, only syntactic criteria are admitted. Identification of basic order thus depends on what syntactic notions the theory attempts to account for, and on assumptions regarding the appropriate means for doing so (cf. McCawley 1970 on English, Stucky 1981 on Maku). In what follows, I will not directly concern myself with this tradition; however, languages like Papago should lead us to consider the extent to which underlying structure can be simply unordered.

Claims regarding basic order in the closely related dialects Pima and Papago (Uto-Aztecan) have varied a great deal, depending partly on whether one follows the standard typological tradition or the syntactic tradition of underlying order. Hale 1975 identifies Papago basic order as SuOV on the basis of intonation.² Later (1983:302), he posits phrase structure rules which 'indicate that non-clausal complements [noun phrases] are prenuclear'. Here Hale is concerned with underlying order, and notes that this order 'is not always true at surface structure ... We may assume that Papago makes use of extraposition to derive alternative word orderings.' Saxton & Saxton 1969 and Langacker 1977 identify Papago as VSuO; Saxton 1980 describes it as predicate non-final; and Saxton 1982 describes it as predicate-initial. Munro 1984 leaves the question open for Pima, noting that all possible orders occur in elicited materials.³ On the supposition that VSuO is basic (presumably within the standard typological tradition), Hawkins (1979, 1980, 1983) lists Papago as VSuO. As will become evident, the criteria of Mallinson & Blake point to clauses of the VSuO or VOSu types as basic. Alternatively, if subjects should be definite and objects indefinite in the type of clause used for determining basic order, OVSu would be basic for Papago.

These conflicting possibilities point out the need to address several issues in addition to that of typological vs. syntactic underlying order. Let us briefly review several problems relevant to studies within the standard typological tradition. First, for some languages, the notions of subject and/or object are difficult to define in a way which insightfully reflects the facts of those languages (cf. Schachter 1976 on Philippine languages, Li & Thompson 1981 on Mandarin Chinese, Schachter 1984 on Toba Batak). There are also languages with ergative case-marking or morphology, and others with agent/patient marking, where the relevance of subject and object categories for constituent order has not always been substantiated.

In a great number of the world's languages, however, identification of subject and object is relatively unproblematic. Yet difficulties may remain in identifying basic order; e.g., one fixed order may occur in a particular environment, but another elsewhere. This is the well-known claim for standard German, where subordinate clauses have SuOV order, while main clauses have VSuO order. Shaul (1986:130) argues that, in colonial Tepehuan (Uto-Aztecan), main clauses

² Hereafter I use the abbreviations SuOV, SuVO, and VSuO, etc. in order to distinguish subject (Su) from the single argument of an intransitive clause (S).

³ Anderson 1976 argues for SuVO in Pima, based on the distribution of certain particles. Drawing on elicited Pima sentence material, Pamela Munro (p.c.) suggests that VOSu may be preferred over VSuO.

manifested a verb-initial pattern with some freedom of order, while subordinate clauses had stable SuOV order. In other languages, order may be so 'free' that one must include discourse considerations. But even discourse studies may reveal no preferred order of subject, object, and obliques relative to one another and to the verb—since order in such languages is often dependent on the pragmatic status of the information encoded (cf. Blake 1983 on Kalkatungu).

Despite these substantive problems, the landmark work of Greenberg 1966, as well as Hawkins' extensions of it, suggest that the dominant order of major syntactic roles, coupled with order of elements in noun and adpositional phrases, forms a significant typological base for a large subset of the world's languages. Nevertheless, the notion of basic order as applied within the standard typological tradition has often been somewhat intuitive. The sources from which typologists have drawn do not apply consistent criteria. In some cases, too, the original claims are not based on accurate research which includes natural discourse.

All these problems ask for clarification of how the term 'basic', in the standard typological tradition, should be identified. One recent proposal suggests that basic order should be determined by that which occurs in main, active, declarative clauses used in contexts in which the subject is DEFINITE and easily identified, and in which the object is INDEFINITE but REFERENTIAL (cf. Givón's 1979 discussion of English).⁴ Mallinson & Blake propose that basic order be determined by that which occurs in stylistically neutral, indicative clauses with definite direct arguments expressed by full NP's. Presumably such proposals refer to the most frequent order within such a clause type. These two proposals simply push the question back one step—forcing us to ask whether introducing indefinite (generally new) information, or talking about definite (generally already given) information, is a more basic communicative function. Both are clearly essential to communication.⁵

When we are faced with naturally occurring oral discourse, these proposals still leave us with a major problem. It can be difficult to find a sample of the stipulated type of data which is large enough to draw any valid conclusions about 'basic order'. In natural discourse, transitive clauses rarely have two overt (non-pronominal) NP's; and it is particularly rare to find such clauses where the clause is not pragmatically marked and both overt NP's are definite (Mallinson & Blake's criteria). Instead, the communicatively basic clause-type has one or zero overt arguments, whether it is transitive or intransitive (cf. D. Payne 1985 for one study; see Chafe 1980 for possible explanations).

Here I examine order in Papago from a discourse perspective. Surprisingly, the text data show that—even in the absence of case-marking—no strongly preferred order of subject, object, and verb can be identified. However, there is a strong order condition on types of information (§5), and semantic factors play a role in the choice of syntactic role (§6).

⁴ Definitions of definite, indefinite, referential etc. are given in §2, below.

⁵ Hawkins (1983:13) makes an additional proposal which, while allowing that discourse-pragmatic factors sometimes motivate order variation, is less explicit about which informational statuses should be taken as unmarked.

2. **RELEVANT DISCOURSE-PRAGMATIC CONCEPTS.** Five discourse-pragmatic concepts prove important in accounting for the data. For the first three, information may actually have one of several values along a parameter. For the fifth, information may occur in one of several specific subcontexts.

IDENTIFIABILITY: Non-identifiable, or INDEFINITE, information is that for which the speaker assumes that the hearer will not be able to pick out and establish reference based on information already available within the universe of discourse (cf. Chafe 1976, Du Bois 1980). When introducing information which is indefinite and referential, speakers are most commonly requesting their hearers to open an ACTIVE DISCOURSE FILE for it; afterwards, it is treated as identifiable.

IDENTIFIABLE, or DEFINITE, information is that for which the speaker assumes that the hearer can pick out and establish reference, given information already available within the universe of discourse. I use the term NON-INDEFINITE to refer to the conjunction of definite and non-referential information. This proves to be a significant language-specific category for information ordering in Papago.⁶

REFERENTIALITY: Referential entities or concepts are those which are treated as existing, bounded entities within the universe of discourse. They have continuous identity over time (Du Bois 1980). Information may, alternatively, be non-referential. Information which is semantically indefinite and referential—but which, pragmatically, is supremely unimportant—is sometimes treated as if it were non-referential (cf. Givón 1985).

GIVENNESS: Given information is that which the speaker assumes to be immediately accessible within the hearer's active consciousness. (In reality, there are various degrees of accessibility; cf. Chafe 1976, 1984.) Other information is new: it may not have previously been in the hearer's memory store at all, or it may be in long-term memory, or in the peripheral consciousness.

An ACTIVE DISCOURSE FILE refers to a mental entry that is created, 'lit up', or activated for an entity or concept within a person's (here, the hearer's) active focus of consciousness (Chafe 1984). Once a file is activated for an entity, this entity becomes available for future deployment as a participant or prop within the discourse; and it can be referred to as the same entity, often by anaphoric devices. It is then normally treated as definite and referential within the universe of discourse. Once the entity fades into background awareness or peripheral consciousness (because of non-mention for a time, or discourse discontinuity, or some other factor), then a new cognitive file need not be established in order for it to be re-mentioned. However, some discourse-pragmatic device is often employed in order to re-activate an already existing file. For certain entities and concepts, such as culturally unique items, categories are permanently available.

⁶ Sandra Thompson has suggested the term '(Non-)identifiable' for the informational status and '(non-)identifiable' for a formal morphosyntactic marking. This nicely clarifies the terminology. However, in this study it would result in some awkward phraseology such as 'non-non-identifiable'. Here '(non-)identifiable' and '(in)definite' are used synonymously.

PRAGMATICALLY MARKED is a cover term for information which is pragmatically non-neutral, relative to its communicative force. Specifically included are information in single-focus and double-focus contrast situations (Chafe 1976, Dik et al. 1981); information questions and answers to them; and other situations or constructions which have similar contrastive make-up (cf. Dik et al.). New information is not necessarily pragmatically marked, since it need not be contrastive. A third type of pragmatically marked information is contrastive change in MAJOR TOPIC: this is the participant that a text or subtext is about (Givón 1983 suggests objective measures for quantifying the degree to which an entity is a major topic).

3. **BASIC CLAUSE STRUCTURE.** As further background, I outline here a few details of Papago/Prima clause structure. All possible orders of subject, object, oblique (adpositional) phrases, and verb occur.⁷ Exx. 1–2 illustrate several combinations (order within adpositional and genitive phrases will not concern us here):⁸

- (1) a. Su aux V O obl.
 'uvi 'ai ha-hidol heg cucul hidolidakud 'ed.
 woman pf 3pl-cook DET chickens pot in
- b. obl.
 Hidolidakud 'ed 'ai ha-hidol heg cucul heg 'uvi.
 'The woman cooked the chickens in the/a pot.'

- (2) a. O aux V
 'uvi 'an-i valda.
 woman lsc-pf dance, with
- b. V aux Su
 Valda 'an-i heg 'uvi.
 'I danced with the/a woman.'

There is no nominal case-marking in Papago/Prima. As these examples show, a determiner *heg* (Papago *g*) or a non-reduced demonstrative precedes an NP when it is in non-initial position. Distribution of the (reduced) determiner is strictly syntactic, identifying the beginning of non-initial NP's (excluding free pronouns and various complex phrase-initial NP's). Its use does not correspond to categories commonly associated with determiners with 'identifiable' or 'non-identifiable' senses.

⁷ In §4.1, I discuss subject and object identification.

⁸ Examples 1–3 are Prima from Mrs. Eithelene Romero. Except where indicated, Papago examples are from Saxton & Saxton 1973, identified as S&S plus a page number. Unless specified otherwise, the orthography used here is an adaptation of details discussed in the same for both dialects. The orthography used here is an adaptation of details developed for Papago by Albert Alvarez, Ken Hale, and others.

Abbreviations are as follows: A = the most agentive argument of a two-argument clause (see that developed for Papago by Albert Alvarez, Ken Hale, and others).
fn. 1): aff = affirmative; AM = the particle 'am', which sometimes has locative/directional functions; aux = auxiliary; COND = conditional; DET = determiner; DS = different subject; FUT = future; GEN = genitive; ITR = initiator; LOC = locative; MOD = modal; NEG = negative; O = the non-A argument of a two-argument clause, or object; obl. = oblique; PCT = punctual; PF = per-fective; PL = plural; OR = quotative mood; REF = reflexive; REM = remote; S = the single argument of a one-argument clause; SO = singular; SS = same subject; SU = subject; V = verb.

Most clauses contain what Uto-Aztecanists have termed an 'auxiliary' (AUX):⁹ this almost always occurs in second position following the first constituent of the clause (in Papago, it can be in first position.) The AUX indicates aspect, mood, and person/number of subject. There is no overt marking for 3rd person subject, though the presence of an AUX without person/number marking in any non-imperative sentence may be taken as indicating 3rd person. If the AUX is absent, the referent could be of any person/number combination. As illustrated in ex. 1, a set of verbal prefixes indicates person and number of the object. Ex. 2 shows that 3sg. objects are not marked on the verb.

In connected discourse, clauses are often introduced with *c*, *k(c)*, or *ku*; these morphemes function like a switch-reference system (Hale 1983). Alternatively, clauses may begin with an initiator or introducer (itr) *m*, *n*, or *b*. These elements count as the first constituent of the clause, and AUX elements may be suffixed to them:

- (3) a. *M-an-t vaila heg 'uvi.*
itr-1sg-PF dance.with DET woman
b. *Ku-n-t vaila heg 'uvi.*
DS-1sg-PF
'I danced with the woman.'

It is not necessary that the verb be contiguous to the AUX. One or more NP's, postpositional phrases, adverbials, locatives, directionals, or other grammatical elements can intervene between the AUX and the verb:

- (4) *ʃ 'am 'ep cum babiugacud.*
QT AM again MOD making.necklaces.for
'Again he tried to put it around (her) neck.' (S&S 13)
(5) *K waʃ-kiaP meek ga huu haha waʃ 'i hihim ...*
SS still far over.there REM just PCT going.PL
'While they were still a long way off ...' (S&S 13)

4. FACTORS MOTIVATING CONSTITUENT ORDER. Order cannot be accounted for directly by semantic case role in Papago/Pima, since agents can precede patients and vice versa. Further, assignment of nominals to subject and object status cannot be done simply on the basis of semantic role. Benefactives and recipients, as well as patients, can be encoded as direct objects. Patients can be encoded as subjects in reflexive passive clauses (Saxton 1982:143, 144, 225). In what follows, I will consider textual evidence for syntactically and pragmatically determined order.

4.1. METHODOLOGICAL PRELIMINARIES. Since results may depend crucially on methodological coding decisions, I will be as explicit as possible about those followed here. Though these decisions are specific to Papago, similar issues must be addressed in all studies of this sort.

This study is based on a corpus of thirteen Papago and two Pima texts. Nine of the Papago texts are taken from S&S; they consist of legends, stories, or

⁹ This is not a verb, but a second-position complex indicating person, number, aspect, and mood. Conditions under which the AUX is absent do not concern us here.

event sequences which are well known in the culture (two might be termed 'procedural', while seven are 'narrative'). These texts, which were studied in detail, contain a total of 759 usable clauses. Copular clauses and what appear to be false starts were not included. Two other Papago texts, consisting of 145 usable clauses, were provided by Dean Saxton. These are first-person narratives with a large percentage of third-person NP's which are not well known in the culture. These two texts were compared with the S&S texts primarily with regard to order phenomena.¹⁰

For this study, a clause is identified as a predication which contains a single non-nominalized verb. Division into clauses is aided by the switch-reference morphemes and AUX's, such that no more than one of each class is considered to be within a single clause. Relative and other subordinate clauses are counted as separate tokens from their matrix clauses, though head plus relative clause is counted as one argument of the matrix clause. There are no obvious order differences between main and subordinate (or embedded) clauses; indeed, the distinction between main and subordinate is not always clear.

Both subject and object are termed DIRECT arguments of the verb. The subject is that nominal which is referenced in the AUX; the object is identified as any other NP which is not part of an oblique or postpositional phrase. Person/number of object are referenced in the verbal prefix when not 3sg. When subject and object are both 3sg. or both 3pl., there is potential ambiguity regarding subject vs. object, since there is no overt marking in these cases—nor is there nominal case marking. Again, when the AUX is missing and the object is 3sg., both subject and object are referenced by zeroes. In practice, ambiguity rarely arises; when it does, the switch-reference system can help disambiguate.

When no auxiliary occurs within the clause, the subject is that nominal which is co-referential with the subject of the preceding clause. Both direct and indirect object are counted simply as objects; e.g.,

- (6) *K maak g ceoj g gogs g cuukhug.*
SS give DET man DET dog DET meat
'And the man is giving the dog meat.'

(Saxton 1982:108)
Neither *m*-initial complement clauses nor extended complements of verbs of

¹⁰ Two Pima narratives were also compared with the S&S texts, but no significant differences in the factors studied here were found. The Pima texts are from Mrs. Ehelene Rosero, transcribed and glossed by Pamela Munro. These are 'The coyote and the jackrabbit' and 'A reminiscence'. The two unpublished Papago personal narratives are: 'A hunting trip' by James Adams, and 'Great Uncle gets lost' by Susie Enos. The S&S texts used are 'Coyote imitates Skunk', 'How to play field hockey', 'Marriage customs', 'A mean ruler is assassinated', 'How some stars appeared', 'How everything began: The Milky Way appears', 'A quail escapes the Hawk', 'Cottontail tricks Coyote', and 'The story of the beast'. These texts have undergone a certain amount of editing, but S&S write (xviii): 'A written literary style has not yet developed in O'otham [= Papago] ... [In these texts] the careful O'otham reader will find many instances of unpolished oral style on which to practice producing literary style.' These texts were broken into morphemes and glossed with the help of Lynn Gordon, Hyo Sang Lee, Pamela Munro, Janine Scanlon, Olivia Tsosie, and Charles Ulrich. Although not included in this study, the oral Papago texts in Bahr et al. 1974 contain a high percentage of non-referential mentions; they would permit an interesting comparative study with the S&S texts.

saying are counted as objects, since they do not behave as objects for purposes of AUX placement (cf. Munro 1982:310).

In floated quantifier constructions, a quantifier precedes the verb; but the NP to which it pertains (if overt) almost always follows the verb. Munro 1984 argues that floated quantifiers form a constituent with the verb. Consequently, rather than count one preverbal and one postverbal subject constituent in constructions like 7, just one postverbal subject NP is counted:

- (7) *ʃ* 'am *hebaɪ* *ha i* 'o'odhamag *g* *kakaɪcu*.
 OT AM somewhere some populate DET qual

'They say there were some qualis living somewhere.' (S&S 142)

In many clauses like 4-5, several adverbials and other grammatical elements precede the verb.¹¹ Counting these clauses as verb-initial may not be a fair reflection of the language, even in the absence of preverbal direct arguments. Furthermore, AUX is a strongly second-position element, and this argues that a clause should be counted as truly verb-initial only when the verb precedes the AUX:

- (8) *Maasi* 'at.
dawn(VERB) PF
 'It's morning.' (S&S 19)

In the texts studied, this is the only instance in which the verb occurs in syntactic initial position. In the tables which follow, clauses listed as VSuO, VSu etc. are verb-initial only in the sense that the verb precedes the nominal arguments.

Information is categorized as to whether, at the time of speaking, it was presumed to be identifiable by the hearer (definite vs. indefinite), and whether it was presumed to be given or new (cf. §2). These judgments mean that some subjectivity is introduced into the analysis, since I cannot be sure what the speaker or editor of the texts actually presumed.¹² Direct and oblique arguments are counted as definite if they are presumed not to be identifiable by the hearer at the time of speaking, but are immediately modified by an identifying clause.¹³

¹¹ Two questions not pursued here are (a) whether all these preverbal elements and the AUX form a single constituent as part of the VP; and (b) the order of noun and postpositional phrases relative to these preverbal elements.

¹² Identifying 'new' in terms of first mention, and 'given' in terms of non-initial mention, is less subjective; however, such definitions prove faulty for two reasons. (The relative unimportance of 'first mention' as a linguistic category was initially suggested to me by John Du Bois.) First, some non-initial mentions precede the verb, and thus behave more like indefinite/new information (§5). Some of these occur after lapses of 15 or 16 clauses after previous mention. This fact suggests that first mention is not a significant linguistic category, but that identifiability is (Chafe 1976). Second, some entities behave upon first mention more as if they were given information (§5). Thus, in one text, a house is introduced in preverbal position (i.e. as new information); subsequently, mention is made of a door of the house, but the door is introduced postverbally (i.e. as definite/given). This further suggests that information invoked by a frame can be presumed to be identifiable, or in the hearer's consciousness (Christopherson 1939): given a house, we can assume in many (but not all) cultures that a door, or the door, to the house is also available in the consciousness.

¹³ 'Specific' may be a more correct designation for these arguments than 'definite'. However, in order to maintain adequate cell sizes, I do not differentiate these labels.

The modifying clause may be a relative clause beginning with *m-* as in 9, or with *c* as in 10:

- (9) *M-an-i* 'ab *wo* *wa* 'i *hū heg* 'eda *taɪ* [m-ap 'am
 TR-1SG-PF LOC MOD expected PCT go DET in day TR-2SG AM
hab 'aagɪ.
 thus tell

'I will come on the day you say.' (S&S 355)

- (10) *M-aɪ* 'imhab *ʃi al* *taagɪ* *daha* *g* 'o'odham
 TR-QT here east toward sitting DET man
 [c *g* 'uvi *s-behɪdag*].
 SS DET woman AFF-getting
 'There was a man in the east (who was good at) getting women.'
 (S&S 17)

Unique entities, like *jewed* 'earth' in 11 and *ʃi al* 'east' in 12, are considered definite—since their uniqueness allows immediate identification, even if they have not been previously mentioned or implied:

- (11) *K* 'i'ajed *cum* *gaag* *gʻik* *taɪ* 'ab *c* *was* *cemo* *g*
 SS from.here MOD looking.for four day in SS just cover DET
jewed, *ʃ* 'am *jiwɪa* 'uuhum.
 earth QT LOC arrive back.to
 'When he had looked for her for four days and covered the earth,
 he came back.' (S&S 17)
- (12) *N-i* 'im *wo* *gei* *ʃi al* *weco* ...
 1SG-PF here FUT fall east under
 'But if I fall in the east ...' (S&S 19)

Determination of uniqueness is sometimes dependent on knowledge of the culture. In 13, the uniquely original *Siwuliki* 'dust devil' is brought into being. That this is the original *Siwuliki* is known to the hearer via knowledge of the folklore and mythology, and by previous information in the text as to who is the creator (Dean Saxton, p.c.):

- (13) *T* 'am *hahawa* *naato* *g* *Siwuliki*.
 PF LOC then make DET dust.devil
 'Then he made the (original) dust devil.' (S&S 15)

Phrases which are neither clearly definite nor indefinite, such as non-referential mentions, are included in an 'other' category. The 'other' category also includes NP's referring to mass entities when their referentiality and identifiability are unclear. Thus, in 14, *juuki* 'rain' and *ha'icu* 'food' are mentioned for the first time in a text:

- (14) 'I'ioi 'aɪ 'am *wuuppa* *g* *juuki* *c* *wes* *ha-gegosid* *hegam*
 (name) OT AM lay.down DET rain SS all 3PL-fed those
 'o'odham. 'I'ioi 'aɪ 'am 'e'aga *g* *ha'icu* *hugi*.
 people (name) OT AM plant DET something food
 'I'ioi made rain fall and fed all the people. I'ioi planted food.'
 (S&S 349)

The last clause here does not mean 'I'ioi planted some food': *ha'icu hugi* is the lexicalized expression for 'food'. *Juuki* and *ha'icu hugi* are not clearly bounded concepts, and may be best thought of as non-referential.

In this study, all free pronouns (including demonstratives which have no accompanying NP) are counted as pragmatically marked, since participants can be adequately tracked in discourse without resort to such a strong device as a free pronoun. In nearly all cases, free pronouns are in fact contrastive.

4.2. SYNTACTIC ROLE ORDERS. In the following discussion, 'zero overt arguments' designates clauses which do not refer to any of their direct (non-oblique) arguments with an overt NP or free pronoun, regardless of whether the clause is transitive or intransitive. Similarly, 'one overt argument' refers to clauses with one such NP or free pronoun, and 'two overt arguments' refers to clauses with two such phrases. Table 1 summarizes the percentages of clauses in the S&S texts with zero, one, and two (infrequently three) overt direct arguments. Table 2 summarizes parallel percentages in the two Papago personal narratives. In approximately 32% of the clauses which have one overt direct argument in S&S, the single direct argument precedes the verb. The argument follows the verb in approximately 68% of the clauses. Of the zero direct argument clauses, 83% are intransitive and 17% are transitive. In 11% of instances, they have one or more oblique arguments (pospositional phrase, locative, or time word).

	NO. OF CLAUSES	PERCENT
Zero direct arguments	429	56%
One direct argument	302	40%
Two or more direct arguments	28	4%
Total	759	100%

TABLE 1.

	NO. OF CLAUSES	PERCENT
Zero direct arguments	111	77%
One direct argument	31	21%
Two or more direct arguments	3	2%
Total	145	100%

TABLE 2.

It has sometimes been hypothesized (a) that, in culturally familiar stories, such as folkloric narratives, fewer NP's will occur than in culturally novel stories, such as personal narratives; and (b) that those NP's that do occur in the former will contain primarily definite and given information, since so much is assumed to be shared knowledge. However, in the Papago personal narratives vs. the well-known S&S stories (see §4.1), the opposite of (a) is found: both types of text show nearly identical results with regard to (b). The first issue is addressed here, and the second in §5.

Tables 1-2 show that 44% of the clauses from the well-known stories have overt direct NP's, in comparison with only 23% from the personal narratives. The difference is statistically significant, but the reasons await further study. Many factors can affect the percentage of NP's used—such as the number of

semantically transitive vs. intransitive events reported, the number of participants 'on stage' that need to be disambiguated, and (conceivably) language-specific factors such as richness of verbal agreement, clitic systems etc. Presumably, numerous NP's are sometimes used, even in familiar stories, because a story is worth telling only if some information is presented AS IF it were new and indefinite.

As seen in Tables 1-2, clauses with no overt direct arguments are preferred over clauses with one. Both these types are much preferred over clauses with two overt direct arguments. Table 3 shows that, for both transitive and intransitive clauses, the tendency is to have one less overt direct argument than the number theoretically allowed (this pattern is only weakly significant, however). Thus the over-all preference for zero overt direct arguments, seen in Table 1, is partly the consequence of the higher proportion of intransitive clauses in the sample.¹⁴

	TRANSITIVE	INTRANSITIVE	TOTAL
	(2 allowable)	(1 allowable)	
Number of missing arguments	73	31%	73
2	141	60%	498
1	22	9%	188
0	236	100%	759
Total			

TABLE 3.

Table 4 presents syntactic role orders, relative to the verb, for clauses that have two overt direct arguments in S&S (I ignore here the position of preverbal grammatical elements, adverbials, and AUX). Percentages are not calculated, since the total number of such clauses is so small.¹⁵

SuOV	1
OSuV	—
SuVO	8
OVSu	5
VSuO	3
VOSu	3
OVO	4
VOO	2
Total	28

TABLE 4.

As shown in Tables 1-2, the most common clause types in discourse have zero and one overt direct arguments. One-argument clauses say little, of course, about relative order of nominal arguments; however, order of A, S, and O relative to the verb is certain. We might be able to extrapolate from this as to

¹⁴ I ignore here the fact that some transitive clauses may have two objects. This accounts for a seeming discrepancy between Table 1, which shows 28 clauses with two or more arguments, and Table 3, which shows 22 transitive clauses with overt subject and object.

¹⁵ In the Papago personal narratives, only three clauses out of 145 have two overt direct arguments. These have the orders SuOV, OVSu, and VOO.

whether, within the typological tradition, SuOV/OSuV, SuVO, OV/Su, or VSuO/OSu is more basic. Table 5 presents cross-tabulations of syntactic roles (A, S, and O) with pre- vs. postverbal position for clauses with one and two overt direct arguments in S&S. Table 6 presents the same cross-tabulations for the Papago personal narratives.

TABLE 5 ($\chi^2 = 8.46, p < .05; \phi = .15$)

	A	S	O	TOTAL
Preverbal	15	33	44	92
Postverbal	22	136	108	266
Total	37	169	152	358

TABLE 6

	A	S	O	TOTAL
Preverbal	3	6	8	17
Postverbal	1	10	9	20
Total	4	16	17	37

The value of χ^2 for Table 5 is significant.¹⁶ The significance results primarily from the stronger placement in postverbal position of intransitive subjects (S) and objects (O) than would be expected if order were simply random. One factor contributing to the more strongly postverbal placement of intransitive subjects is the presence of floated quantifier constructions. In nearly 100% of these, the NP from which the quantifier is floated (if overt) is in postverbal position.¹⁷ This suggests near grammaticalization of predicate-first order for this construction. In the data, quantifiers were floated only off S's and O's, never off A's.

The χ^2 statistic only indicates whether a particular association is significant. The ϕ statistic, however, measures strength: a value of zero indicates no association, while a value of one indicates an association of 100%. For the data in Table 5, the value of ϕ indicates a very weak association between syntactic role and pre- vs. postverbal order. If just A and O are considered, the distribution in pre- vs. postverbal position is not significant. In §5, I will present data on order of constituents relative to new vs. given, indefinite vs. definite, and pragmatically marked status. Order is almost completely accounted for by these factors.

5. PRAGMATIC FACTORS AFFECTING CONSTITUENT ORDER. The textual data suggest a three-part hypothesis which governs the ordering of information in Papago, relative to the verb:

- (15) a. Non-identifiable (indefinite) information precedes the verb when the hearer is instructed to open a new active discourse file for it, making it available for further deployment.
- b. Pragmatically marked information (including all information question words) precedes the verb.

¹⁶ The cell sizes in Table 6 are too small to allow valid calculation of χ^2 .

¹⁷ This is true even if the NP encodes indefinite information. The relevance of indefinite status to order is discussed in §5, below.

- c. Information follows the verb when the hearer is not instructed to open a new active discourse file for it. This category includes items for which active cognitive files are already available (e.g., identifiable, definite, and unique items)—as well as entities for which files are not to be established, including non-referential mentions.¹⁸

These hypotheses were tested for both direct and oblique phrases. In the S&S corpus of 759 clauses, there are 480 such phrases. Preverbal vs. postverbal position relative to indefinite and definite statuses is shown in Table 7. These data show that the hypotheses in 15 are supported as strong tendencies.

	PREVERBAL	POSTVERBAL	TOTAL
Indefinite	125 83%	26 17%	151 100%
Definite	6 2%	278 98%	284 100%
Pragmatically marked	38 93%	2 5%	40 100%
Other	1	4	5
Total	170	310	480

TABLE 7 (χ^2 with Yates' correction, calculated over indefinite and definite information, = 304.6, $p < .001; \phi = .837$)

The value of χ^2 with Yates' correction, calculated over the 435 tokens of non-pragmatically marked definite and indefinite mentions, is significant. Thus the null hypothesis that there is no association between definite/indefinite status and pre- vs. postverbal order can be rejected. The value of ϕ for the definite/indefinite data shows that the association is very strong indeed.

Table 8 presents similar data for the Papago personal narratives. The value of χ^2 with Yates' correction, calculated over just the 39 tokens of non-pragmatically marked definite/indefinite information, is also strong. The value of ϕ indicates that the strength of the association is very strong.

	PREVERBAL	POSTVERBAL	TOTAL
Indefinite	11 73%	4 27%	15 100%
Definite	1 4%	23 96%	24 100%
Pragmatically marked	6 100%	6 100%	12 100%
Other	16 62%	10 38%	26 100%
Total	34	37	71

TABLE 8 (χ^2 with Yates' correction, calculated over indefinite and definite information, = 17.6, $p < .001; \phi = .731$)

Comparison of Tables 7-8 with Tables 5-6 argues that, both in well-known narratives and in non-shared personal narratives, order is far more sensitive to pragmatic status than to syntactic role.

5.1. THE FUNCTION OF POSTVERBAL POSITION. Although the number of clearly non-referential mentions in the 'other' category is too small for us to say much with certainty, it is clear that non-referential information can follow the verb.

¹⁸ Both 15a and 15b may be subcases of a single cognitive status, concerned with degree of expectancies (cf. Givón 1985). The preverbal information described in 15a is inherently less expected than the postverbal information described in 15c. In 15b, the NEW/ALREADY asserted between the marked information and the rest of the predication may not be expected. A thorough test of this hypothesis would involve psycholinguistic experimentation.

This suggests that postverbal position may not correlate so much with definiteness as with non-indefiniteness, i.e. the conjunction of non-pragmatically marked categories which are not indefinite and referential. However, additional evidence argues for a slightly different analysis. As Table 7 shows for the S&S texts, 17% of indefinites occur postverbally. Nine of the 26 mentions have preverbal floated quantifiers, and it appears that postverbal position is largely grammaticized for NP's with such quantifiers. All four of the postverbal indefinites in Table 8 have these.

Thirteen of the 26 indefinite mentions in Table 7 contain relatively non-salient information which does not have continuity throughout the discourse. These include some time expressions (as in 16), locatives, or ancillary equipment for the action (as in *uʒabi-kaj* 'with pitch', 17):

- (16) T 'am 'i 'e'ai g g'i'ik taʒ ...
 PF AM PCT REF-teach.time DET four day
 'When four days had passed ...' (S&S 356)

- (17) Ku-t-ʒ 'am si bibidʒ g toobi g wuupui-ʒ g
 DS-PF-QT AM INTENSITY plaster DET cottontail DET eyes-GEN DET
 ban g' uʒabi-kaj. Ku-t-ʒ 'ab hahawa mehi cause.to.burn
 coyote DET pitch-with DS-PF-QT LOC then
 g toobi g waapk.
 DET cottontail DET cane
 'Cottontail plastered Coyote's eyes with pitch. Then Cottontail set fire to the cane.' (S&S 127)

These postverbal initial mentions suggest that the function of postverbal position is to direct the hearer NOT to establish a 'new active discourse file' for the entity mentioned—either because such a file already exists, or because the will not be needed. Although the item may momentarily enter the active consciousness, the hearer is not instructed to 'set up' a cognitive file for anaphoric purposes.¹⁹ The latter function is perhaps identical to Givón's identification (1984) of semantically referential entities which, for all pragmatic purposes, are non-referential and are treated as such (cf. also Du Bois 1980, Schuh 1977). Pragmatically, these entities are supremely unimportant. Hypothesis 15a is supported by 97% of the data from S&S when clauses with floated quantifiers and with unimportant indefinites are factored out (i.e., 125 out of 129 apparently non-incidental referential indefinites which also do not occur with floated quantifiers). Hypothesis 15a is supported by 100% of the data from the personal narratives under similar conditions.

5.2. PRAGMATICALLY MARKED INFORMATION. The distribution of types of pragmatically marked information is shown in Table 9. Although some comments are warranted here.

¹⁹ Nevertheless, items which fail to have discourse continuity can occur preverbally. The hypothesis is that active discourse files are opened for such items, and that they are potentially available for future deployment.

S&S texts:	5
Single focus contrast	4
Double focus contrast	7
Fronting of contrastive topic	24
Free pronoun	40
Total	
Papago personal narratives:	3
Double focus contrast	3
Free pronoun	6
Total	

TABLE 9.

A 'highly topical' participant is that one which a text or subtext is about. This participant generally has continuity throughout some portion of the text (Givón 1983). In the corpus studied, fronting of a contrastive topic is generally restricted to given information which was last mentioned within twelve clauses. It is contrastive in the sense that 'I as speaker am no longer going to be talking about X as topic, but Y.' It is similar to the topic-establishing function of the *for X ...* phrases in English. Fronting of a contrastive topic is NOT equivalent to 're-introduction' or re-activation of information in the consciousness of the hearer after a significant lapse. Lapses of the latter sort center around a distance of twelve to sixteen clauses in the texts studied, and simple re-introduction is not necessarily contrastive. In 18, six clauses intervene between (b) and (d), discussing how women decorate hockey sticks. The word 'ola' 'hockey puck' is given information (but non-topic) in line (b); however, line (d) begins a section in which the puck is highly topical, describing how it is made. This correlates with fronting of 'ida' 'ola' 'this puck':

- (18) a. ma-i heg hekaj wo s'-ap 'i 'e-naaʒ 'an 'e-kuug 'an
 ITR-PF that with MOD AFF-good PCT REF-fold LOC REF-end LOC
 '... so that it can be bent just right at the end'
 b. ma-i hekaj s'-ap wo wuic g 'ola ...
 ITR-PF with AFF-good MOD cause.to.come.to DET puck
 'for tossing the puck ...' [five clauses intervene]
 c. T hab masma 'an wo 'e-o'oha g ha-'usaga.
 PF thus like LOC MOD REF-draw DET 3PL-hockey.stick
 'That is how the hockey stick is decorated.'
 d. K 'ida 'ola hab masma 'e-naaʒ ...
 SS this puck thus like REF-finish
 'The puck is made ...' (S&S 219-20)

In most cases, free pronouns occur in contexts of single or double focus contrast (Chafe 1976); however, some cannot be accounted by contrastiveness. Dik et al. (1981:42) formally define six cleft focus subtypes, some of which need not be contrastive. This richer inventory accounts for a good portion of the residual cases. The two postverbal instances of (presumably) pragmatically marked information in Table 7 are 3rd person pronouns.

Finally, the grammaticized position for information question words is pre-

verbal, regardless of whether they are used in questions or as indefinite pronouns:

- (19) K *pi hedai 'ia huu i-tai-cua.*
 SS NEG who here REM 1PL-want

No one wanted us here. (S&S 11)

Phrases which provide information requested also occur preverbally. The corpus contains few answers to information questions, and all these refer to indefinites. Such phrases are counted in the indefinite preverbal category. However, information questions and answers to them are similar to focus constructions in their pragmatic make-up (cf. Dik 1980), and they could be considered pragmatically marked constructions.²⁰

5.3. GIVEN VS. NEW INFORMATION is not as predictive of order as is identifiability. Table 10 presents pre- vs. postverbal distribution of direct and oblique phrases, relative to new vs. given status in S&S. If we factor out 36 cases of information which is pragmatically marked and given from Table 7, we have a total of 253 phrases which are non-pragmatically marked and given.

	PREVERBAL	POSTVERBAL	TOTAL
New	132 69%	59 31%	191 100%
Given	2 1%	251 99%	253 100%
Total	134	310	444

TABLE 10 (χ^2 with Yates' correction, = 238.2,
 $p < .001$; $\phi = .737$)

The value of χ^2 with Yates' correction for Table 10 is significant. The value of ϕ is .737, showing a fairly strong association of given/new status with order. But since the value of ϕ for definiteness/indefiniteness is .837 (Table 7), the latter parameter remains a stronger predictor of order. The percentages in Tables 7 and 10 suggest that postverbal order can be equally well predicted on the basis of either given or definite status. This is to be expected: if information is assumed to be already in the hearer's active focus of consciousness (i.e. given), then we assume that, in nearly all cases, the hearer has established the identity of that entity (i.e., it is definite). The discrepancy enters with new information. In the majority of cases, new information may have unique identity. But a significant percentage can be assumed given a certain context; or they are identified simultaneously with the act of introduction. In sum, the high correlation between given/new status and order falls out as a result of the way in which given/new is associated with identifiability.

As we have seen, however, identifiability does not tell the whole story. What is truly predictive in Papago is whether information is (a) pragmatically marked and (b) whether it is (potentially) important enough for the speaker to instruct

²⁰ Saxton (1982:112-18) states that 'topicalized' phrases occur before the predicate. These convey new or contrastive information, ask for information, or supply responses to information questions.

the hearer to mentally 'tag' the entity as something to be available for further deployment (and which does not occur in a floated quantifier construction). These two categories occur preverbally. Everything else is lumped together in postverbal position.

6. PRAGMATIC FACTORS AFFECTING CHOICE OF SYNTACTIC ROLE. Previous sections have explored the structuring of information relative to noun phrase occurrence, and order of noun phrases vis-à-vis the verb. We now turn to structuring relative to syntactic encoding as A, S, or O. For Papago, the conjunction of A and S can be taken as subject, and O as object. In addition, I use the term 'absolutive' to refer to the conjunction of S and O when they function in opposition to the 'ergative' category A.²¹ The data below add to a growing body of evidence that subject and object—or even A, S, and O—cannot be taken as ultimate primes for a theory of language. The very existence of such categories results from communicative needs. In Papago, animate vs. inanimate correlates with encoding as subject (S, A) vs. Oblique. Besides, the need to distinguish continuity of identity from new informational status motivates two distinct discourse categories: the conjunction of A and S (subject) vs. that of O and S (absolutive). The reality of such principles for structuring discourse becomes more apparent when we find similar patterns operating across unrelated languages. For example, based heavily on Saca-pulce (Mayan) data, Du Bois 1985 argues for identical correlations between continuity of identity and the subject category, vs. new information and the absolutive category. Scancarrelli 1985 discusses similar patterns for Chamorro. In addition, however, I will suggest that O (along with oblique) is more strongly associated cross-linguistically with discontinuity.

Of the clauses with one overt direct argument in S&S (Table 1), 79% of the arguments are A's, 55% are S's, and 38% are O's. This shows that the argument is S for intransitive clauses, and tends to be O for transitives. In 93% of all clauses, then, the overt NP reflects the absolutive category. Insofar as V plus arguments are A's, S's, and O's, and tends to be O for transitives, in 93% of all clauses, then, the overt NP reflects the absolutive category. Insofar as V plus arguments are A's, S's, and O's, and tends to be O for transitives, in 93% of all clauses, then, the overt NP reflects the absolutive category. Insofar as V plus arguments are A's, S's, and O's, and tends to be O for transitives, in 93% of all clauses, then, the overt NP reflects the absolutive category.

Du Bois suggests that S and O are preferred over-all because they are preferred for introducing new information. The Papago data in Table 11 partially confirm this, but also present other details. New information is coded as subjects by the absolutive category: 30% of new phrases are coded as subjects (the sum of S and A), while 60% are introduced as absolutives (N = 97), fully 90% coded by the absolutive category. However, Table 11 (overleaf) also shows that S and O are the preferred roles for all overt direct phrases, and that the majority of overt S's and O's actually encode given information.

²¹ This use of 'absolutive' is not to be confused with the Uto-Aztecanist use of 'absolutive' for certain noun endings.

	GIVEN	NEW	TOTAL
A	27	10	37
S	136	33	169
O	98	54	152
Oblique	73	49	122
TOTAL	334	146	480

TABLE 11 ($\chi^2 = 17.06$, $p < .01$; $\phi = .189$)

This table further shows that more new information is encoded as O (or oblique) than as S. This does not mean that the major discourse function of O is to signal discontinuity, but simply that it is more strongly associated with discontinuity than is either S or A.²² In part, this may be a consequence of its opposition to the grouping of S and A, which is highly associated with encoding continuous/thematic participants (Givón 1983).

Distribution of indefinite information is similar to that of new information. Table 12 shows that indefinites are most commonly encoded as O's or obliques (cf. Givón 1979:52, with similar conclusions for English). Within the subject category, 79% of indefinite mentions are S's, and 21% are A's. This parallels an over-all ratio of 82% S's to 18% A's for overt subject phrases in the data.²³

	DEFINITE	INDEFINITE	TOTAL
A	31	6	37
S	147	22	169
O	90	59	149
Oblique	56	64	120
Total	324	151	475

TABLE 12 ($\chi^2 = 61.6$, $p < .001$; $\phi = .36$)

The value of χ^2 for the data of Table 11 is significant. The value of ϕ is a relatively low .189. The value of χ^2 for the data of Table 12 is also significant, but the value of ϕ is .36. From these differences, we can argue that identifiability is more important in determining syntactic role than is given/new status.

Tables 11-12 indicate that the subject most dominantly codes continuous, topical participants. Of all subjects, 79% encode given information; 86% encode definite information. Of all given subjects, 83% are S's, and 83% of all definite subjects are S's. These percentages show that S plays a strong role in encoding continuity. With regard to overt objects, 64% are given and 60% are definite. Obliques are 60% given and 47% definite. Given that S is strongly associated with continuous definite information, and given that O is not clearly more

²² The articles in Givón 1983 argue that O is more strongly associated with discontinuity than is subject (Su) in numerous languages; however, they do not always differentiate S from A. Switch-reference and continuity systems (cf. Carlson 1984) offer independent evidence for a relationship between O's and discontinuity: S and A function together as a separate category from O, with same-subject chains conveying continuity. When a participant switches from object to subject status, or vice versa, then discontinuity occurs. Further evidence comes from some Amazonian languages which use object clitics to encode S's at points of significant locational discontinuity in narrative discourse (T. Payne 1984).

²³ Non-referential 'other' phrases are factored out of Table 12 (compare Table 7), since the identifiability contrast is not relevant for non-referential mentions.

strongly associated with new indefinite information than with continuous definite information, questions arise as to the strength and universality of a correlation of new information with the absolute.²⁴

Since new information can be introduced as S, A, O, or oblique, it is of interest to know what conditions precipitate which choice. The hypothesis was made that choice of syntactic role would vary with animacy. Table 13 presents cross-tabulations of the various roles with animacy. In this sample, new inanimates are never coded as A, and almost never as S. Animates are never coded as obliques, and rarely as O.

	ANIMATE	INANIMATE	TOTAL
A	10	—	10
S	29	4	33
O	11	43	54
Oblique	—	49	49
TOTAL	50	96	146

TABLE 13 (χ^2 with Yates' correction, calculated over indicated cells, = 82.2, $p < .001$; $\phi = .77$)

The total number of new A's is very small: we cannot say with certainty that A's are distributed differently than S's. When A and S (subject) are opposed to O and oblique, the value of χ^2 with Yates' correction is significant. The value of ϕ is .77—showing that new animates are strongly associated with subject, and new inanimates with O/oblique. Thus, although new information is more strongly correlated with O's (and obliques) over-all in the data, this relationship might partly result from introduction of more new inanimate entities than animates. The texts are primarily folkloric, with relatively few major animate participants per text. Note, however, that 20% of all animate participants are still introduced as O.

7. **TYPOLOGICAL CONCLUSIONS.** Cross-linguistically, linear order is found to correlate with syntactic role or with pragmatic status (Thompson 1978).²⁵ Syntactic role is the dominant factor in languages like English (SuVO), Chickasaw (Muskogean, SuOV), and Rapa Nui (Polynesian, VSuO; Chapin 1978). It is important to keep in mind, however, that 'subject' and 'object' may not always be the relevant syntactic categories; cf. Toba Batak. In probably all syntactic order languages, pragmatically marked information may or must be shifted to the front of the clause. But we can still term them syntactic role order languages, since order is not primarily sensitive to statuses of given/new, definite/indefinite, topic/comment etc.

Second, order may be primarily correlated with the pragmatic status of in-

²⁴ Although his conclusions are not based on discourse data, T. Payne 1982 discusses evidence for an association between the absolute category and definiteness/hopicality in Yup'ik Estimo and Tagalog.

²⁵ To Thompson's typology we might add strict correlation of order with semantic role (agent, patient etc.) There are no clearly documented cases of this type, though Dik 1980 argues for some candidates.

formation encoded. Papago order is almost completely accounted for by the principles stated in 15 above. Givón 1985 suggests that the Papago pattern of placing indefinite, new, or less expected information first may be more common than previously expected. Other languages which operate on this general pattern include Iroquoian (Malhotra 1984), Algonkian, and Ute (Givón 1983). Additional pragmatic order languages include Lisu (Hope 1973), Kalkatungu (Blake 1983), and Slavic languages like Czech (Firbas 1964). In languages where order is primarily determined by pragmatic principles, there may yet be minor constructions in which order is more strictly based on syntactic factors; this is true of Papago sentences with floated quantifiers.

Third, some languages may evidence a greater balance of pragmatic and syntactic factors, e.g. Mandarin Chinese (Sun & Givón 1985) and Spanish (Benavogue 1983). In these languages, order of subject is largely determined by pragmatic factors, but the object more consistently comes after the verb, and may be syntactically determined.²⁶

Some theoretical approaches assume that (in most languages) underlying structure is ordered, and that this underlying order is syntactically based. Various principles of a universal or language-specific nature may, in effect, be such that surface order is potentially at variance with underlying order. However, the existence of strongly pragmatic order languages calls us to explore the extent to which underlying structure may not be linearly ordered. If, in a given language, almost all order phenomena are accounted for by pragmatic factors, then why should we postulate an underlying level, an order of syntactic roles which rarely or never occurs? Surface order, by contrast, should be accounted for in some languages by syntactic factors, in others by pragmatic factors, and in most languages by interacting factors of both types.

REFERENCES

- ANDERSON, ALICE. 1976. Pima locationals and related matters. UCLA, M. Baur, DONALD M., et al. 1974. Piman shamanism and staying sickness (K'icim miki-dag). Tucson: University of Arizona Press.
- BENAVOUE, PAOLA. 1983. Topic continuity and discontinuity in discourse: A study of spoken Latin-American Spanish. In Givón 1983:255-311.
- BLAKE, BARRY J. 1983. Structure and word order in Kalkatungu: The anatomy of a flat language. *Australian Journal of Linguistics* 3, 143-75.
- CARLSON, ROBERT. 1984. Narrative connectives in Supit. *Coherence and grounding in discourse*, ed. by Russell Tomlin. To appear, Amsterdam: Benjamins.
- CHAFE, WALLACE. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. Subject and topic, ed. by Charles Li. 25-55. New York: Academic Press.
- . 1980. The deployment of consciousness in the production of a narrative. *The pear stories*, ed. by W. Chafe. 9-50. Norwood, NJ: Ablex.
- . 1984. Cognitive constraints on information flow. *Coherence and grounding in discourse*, ed. by Russell Tomlin. To appear, Amsterdam: Benjamins.
- CHAPIN, PAUL. 1978. Easter Island: A characteristic VSO language. *Syntactic typology*, ed. by Winfried P. Lehmann. 139-68. Austin: University of Texas Press.
- CHRISTOPHERSON, P. 1939. The articles: A study of their theory and use in English. Copenhagen: Munksgaard.
- COMRIE, BERNARD. 1978. *Ergativity. Syntactic typology*, ed. by Winfried P. Lehmann, 329-94. Austin: University of Texas Press.
- DIX, SIMON. 1980. Studies in functional grammar. New York: Academic Press.
- , et al. 1981. On the typology of focus phenomena. *Perspectives on functional grammar*, ed. by Teun Hoekstra et al. 41-74. Dordrecht: Foris.
- DIXON, ROBERT M. W. 1979. *Ergativity*. Lg. 55:59-138.
- DU BOIS, JOHN W. 1980. Beyond definiteness: The trace of identity in discourse. *The pear stories*, ed. by W. Chafe. 203-74. Norwood, NJ: Ablex.
- . 1985. Competing motivations. *Iconicity in syntax*, ed. by John Haiman. 343-65. Amsterdam: Benjamins.
- FIRBAS, JAN. 1964. On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague* 1:267-80.
- GIVÓN, TALMY. 1979. On understanding grammar. New York: Academic Press.
- . 1983 (ed.). Topic continuity in discourse: A quantitative cross-language study. Amsterdam: Benjamins.
- . 1984. The pragmatics of referentiality. *Meaning, form, and use in context* (Georgetown University Round Table 84), ed. by Deborah Shiffrin. 120-38. Washington, DC: Georgetown University Press.
- . 1985. The pragmatics of word order: Predictability, importance and attention. *Language typology and language universals*, ed. by Edith Moravcsik. To appear, Amsterdam: Benjamins.
- GNEVBORG, JOSEPH. 1966. Some universals of grammar, with particular reference to the order of meaningful elements. *Universals of language*, ed. by J. Greenberg. 2nd edn. 73-113. Cambridge, MA: MIT Press.
- HALE, KENNETH. 1975. Papago intonation and word order. Paper presented at the 1975 Uto-Aztecan Conference.
- . 1983. Papago (K'icim). *Uto-Aztecan* 49:299-327.
- HAWKINS, JOHN. 1979. Implicational universals as predictors of word order change. *Linguistics* 16:193-236.
- . 1980. On implicational and distributional universals of word order. *Journal of Linguistics* 16:193-236.
- . 1983. Word order universals. New York: Academic Press.
- HOPE, EDWARD. 1973. Non-syntactic constraints on Lisu noun phrase order. *Foundations of Language* 10:79-109.
- LANGACKER, RONALD. 1977. *Studies in Uto-Aztecan grammar*, 1: An overview of Uto-Aztecan grammar. Dallas: Summer Institute of Linguistics.
- LI, CHARLES, and SANDRA THOMPSON. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley & Los Angeles: University of California Press.
- MILLINSON, GRAHAM, and BARRY BLAKE. 1981. *Language typology*. Amsterdam: North-Holland.
- MCCAWLEY, JAMES. 1970. English as a VSO language. Lg. 46:286-99.
- MITTRUN, MARLUSSE. 1984. Is basic word order universal? Paper presented at the Symposium on Discourse Relations and Cognitive Units, Eugene, Oregon.
- MURKIN, PAVEL. 1982. On the transitivity of 'say' verbs. *Studies in transitivity*, ed. by Paul Hopper and Sandra Thompson (Syntax and semantics, 15), 301-18. New York: Academic Press.
- . 1984. Floating quantifiers in Pima. *The syntax of Native American languages*, ed. by Eung-Do Cook and Donna Gertz (Syntax and semantics, 16), 269-87. New York: Academic Press.
- PAYNE, DONIS L. 1985. *Aspects of the grammar of Yagua: A typological perspective*. UCLA dissertation.
- PAYNE, THOMAS E. 1982. Role and reference related subject properties and ergativity in Yupik Eskimo and Tagalog. *Studies in Language* 6:75-106.

²⁶ Pragmatic factors may govern order of direct relative to oblique/indirect objects in a sizable portion of syntactic role order languages; this is true for Yagua (D. Payne 1985), Susanna Cuming and Pamela Munro have reminded me that it is also true for English and Chickasaw.

- _____. 1984. Split-S and fluid-S marking revisited. *Papers from the Parasession on Lexical Semantics*, ed. by John Richardson et al., 173-84. Chicago: CLS.
- SAXTON, DEAN. 1980. Evidence for basic word order in Papago-Pima. ms.
- _____. 1982. Papago. *Studies in Uto-Aztecan grammar*, 3, ed. by Ronald Langacker, 93-266. Dallas: Summer Institute of Linguistics.
- _____, and LUCILLE SAXTON. 1969. Papago and Pima to English dictionary. Tucson: University of Arizona Press.
- _____. 1973. O'otham hoho'ak a agitha: Legends and lore of the Papago and Pima Indians. Tucson: University of Arizona Press.
- SCANCARELLI, JANNE. 1983. Referential strategies in Chamorro narratives: Preferred clause structure and ergativity. *Studies in Language* 9.335-62.
- SCHACHTER, PAUL. 1976. The subject in Philippine languages: Topic, actor, actor-topic, or none of the above. Subject and topic, ed. by Charles Li, 491-518. New York: Academic Press.
- _____. 1984. Semantic-role-based syntax in Toba Batak. *Studies in the structure of Toba Batak*, ed. by Paul Schachter (UCLA occasional papers in linguistics, 5), 122-49. Los Angeles.
- SCHUH, RUSSELL G. 1977. Bade/Ngizim determiner system. *Afroasiatic Linguistics* 4.101-74.
- SHAUL, DAVID. 1986. Topics in Nevome syntax. (University of California publications in linguistics, 109.) Berkeley & Los Angeles: University of California Press.
- SILVERSTEIN, MICHAEL. 1976. Hierarchy of features and ergativity. *Grammatical categories in Australian languages*, ed. by Robert M. W. Dixon, 112-71. Canberra: Australian Institute of Aboriginal Studies.
- STRUCKY, SUSAN. 1981. Word order variation in Makua: A phrase structure grammar analysis. Urbana: University of Illinois dissertation.
- SUN, C.-F., and TALMY GIVÓN. 1985. On the so-called SOV word order in Mandarin Chinese. *Lg.* 61.329-51.
- THOMPSON, SANDRA. 1978. Modern English from a typological point of view: Some implications of the function of word order. *Linguistische Berichte* 54.19-35.

[Received 18 June 1986;
revision received 25 February 1987;
accepted 11 March 1987.]